

## سیستم پیشنهاد دهنده‌ی منبع مشارکتی مبتنی بر تجزیه تنسور

نگین میثاقیان<sup>(۱)\*</sup>، مهرداد جلالی<sup>(۲)</sup>، محمد حسین معطر<sup>(۳)</sup>

(۱) دانشگاه آزاد اسلامی، واحد مشهد، باشگاه پژوهشگران جوان و نخبگان، مشهد، ایران.

(۲و۳) دانشگاه آزاد اسلامی، واحد مشهد، دانشکده مهندسی، گروه کامپیوتر، مشهد، ایران.

### چکیده:

با گسترش اطلاعات، نیاز به پیشنهاد منابع مورد علاقه کاربران بیش از پیش احساس می‌شود از این‌رو سیستم‌های برچسب‌زنی اجتماعی به‌صورت روزافزون در حال افزایش و متداول شدن می‌باشند. یکی از چالش‌های موجود در چنین سیستم‌هایی ترکیب اطلاعات برچسب‌های انتسابی به منابع با زمان انتساب برچسب توسط کاربران جهت کشف صحیح علایق کاربر با توجه به گذر زمان می‌باشد. در این مقاله، داده‌های دریافت شده از سیستم‌های برچسب‌زنی مشارکتی، توسط تنسور ۴ بعدی جهت نمایش ارتباطات چهارگانه کاربران، برچسب‌ها، منابع و زمان انتساب برچسب‌ها مدل شده است، همچنین با کمک تجزیه و تحلیل چند حالتی، ارتباطات پنهان میان علایق کاربران با زمان کشف می‌شود. از این‌رو تاثیر گذر زمان به‌طور مناسب در ارائه پیشنهاد منبع به کاربر سبب افزایش صحت خروجی پیشنهادات می‌شود. نتایج آزمایشات بر روی پایگاه داده‌ی دنیای واقعی citeulike، بهبود صحت پیشنهادات و کارایی این سیستم را در مقایسه با سایر سیستم‌های موجود نشان می‌دهد.

### واژه‌های کلیدی:

سیستم‌های برچسب‌زنی اجتماعی، سیستم پیشنهاد دهنده مشارکتی، تنسور، تجزیه تنسور

\* عهده‌دار مکاتبات

دانشگاه آزاد اسلامی، واحد مشهد، باشگاه پژوهشگران جوان و نخبگان، مشهد، ایران

پست الکترونیکی: [n.misaghian@imamreza.ac.ir](mailto:n.misaghian@imamreza.ac.ir)

جدیدی برای رده‌بندی محتوای وب است. این لغت از ترکیب دو کلمه فوک<sup>۵</sup> و تکسونومی به معنای مردم و علم رده‌بندی مشتق شده است. فوکسونومی سیستم و ساختاری جدید برای سازماندهی، طبقه‌بندی و تعیین شباهت موضوعی میان منابع از دید کاربران و همچنین کمک نمودن به کاربران و سیستم‌ها برای جستجوی منابع اطلاعاتی در محیط وب است.

در سیستم‌های پیشنهاد دهنده به اطلاعاتی از جمله زمان، مکان، کاربر همکار و... در کنار اطلاعات اصلی زمینه<sup>۶</sup> گفته می‌شود.

زمینه یک مفهوم چند وجهی است که در تحقیقات مختلفی مورد مطالعه قرار گرفته و به چندین روش تعریف شده است [۱، ۲].

مطابق با [۳] زمینه به هر نوع اطلاعی اشاره می‌کند که می‌تواند برای شناسایی موقعیت یک موجودیت به کار رود، به طوری که در مورد سیستم‌های پیشنهاد دهنده، موجودیت می‌تواند یک کاربر، آیتم و یا آزمایشی باشد که توسط کاربر ارزیابی می‌شود [۴]. بنابراین هر اطلاعی راجع به موقعیتی که کاربر آیتمی را مورد آزمایش قرار می‌دهد مثل مکان، زمان، آب و هوا، نوع دستگاه و وضعیت روحی می‌تواند به عنوان زمینه در نظر گرفته شود. سیستم‌های پیشنهاد دهنده‌ای که هر نوع از اطلاعات ذکر شده را به کار می‌گیرند به عنوان سیستم‌های پیشنهاد دهنده‌ی مبتنی بر زمینه شناخته می‌شوند [۱].

اهمیت اطلاعات زمینه‌ای برای پیشنهادات صحیح به طور گسترده در سیستم‌های پیشنهاد دهنده مورد بررسی و استفاده قرار گرفته است [۵، ۶]. الگوریتم‌ها و روش‌های فیلترینگ مشارکتی مرسوم اطلاعات زمینه‌ای را جهت ارائه پیشنهاد به کاربر در نظر نمی‌

گسترش و محبوبیت سیستم‌های برچسب زنی اجتماعی همانند فلیکر<sup>۱</sup>، دلیشز<sup>۲</sup> و... در میان کاربران سبب شده است که محققین زیادی فعالیت‌های خود را به منظور کشف علایق کاربران و رفع نیازهای آن‌ها بر روی سیستم‌های پیشنهاد دهنده‌ی مبتنی بر برچسب معطوف نمایند. چنین سیستم‌هایی در مکان‌هایی که حجم وسیعی از داده‌ها وجود دارد به کاربران در پیدا نمودن بهترین منبع اطلاعاتی کمک می‌نمایند. برچسب واژه‌ای است که توسط یک کاربر به طور آزادانه برای یک منبع مهم انتخاب می‌شود؛ به طوری که فرا داده‌ای راجع به توصیف آیتم نمایش می‌دهد بنابراین می‌تواند به عنوان یک کلید واژه برای شناسایی یا پیدا کردن‌های بعدی یک متن مفید باشد. همچنین برچسب‌ها مکانیزم اصلی استفاده شده جهت جستجوی منابع جدید در سیستم‌های برچسب زنی اجتماعی می‌باشند.

با استفاده از سیستم‌های برچسب زنی اجتماعی، کاربران می‌توانند منابع را در انجمن‌ها به اشتراک گذاشته، بارگذاری کرده و به طور عمده دسته‌بندی کننده‌های شخصی را توصیف کرده و بر روی آن‌ها برچسب‌های خاصی را اعمال کنند. دانش و طبقه‌بندی اطلاعاتی که توسط چنین سیستم‌ها جمع آوری می‌شوند، فوکسونومی نامیده می‌شود. می‌توان گفت اجتماع تمام برچسب

های زده شده توسط یک کاربر پرسونومی<sup>۳</sup> وی را تشکیل می‌دهد در حالی که اجتماع تمام پرسونومی‌ها یک فوکسونومی را تشکیل می‌دهند. در حقیقت فوکسونومی<sup>۴</sup> (رده‌بندی مردمی) عنوان تئوری

<sup>۱</sup> <http://www.flickr.com/>

<sup>۲</sup> <http://delicious.com/>

<sup>۳</sup> Presonomy

<sup>۴</sup> Folksonomy

<sup>۵</sup> Folk

<sup>۶</sup> Context

گیرند. سیستم‌های برچسب‌زنی موجود تاکنون ارتباط همزمان ۲ گانه و یا ۳ گانه بین کاربران، منابع و برچسب‌ها را در نظر گرفته‌اند. با گسترش اطلاعات، نیاز به پیشنهاد منابع مورد علاقه کاربران بیش از پیش احساس می‌شود. در حقیقت علاقه‌ی کاربر در طول زمان ممکن است تغییر پیدا نماید، از این‌رو الگوریتم‌های پیشنهاد دهنده‌ی پایه قادر به درک این تغییرات در طول زمان نیستند. چنین الگوریتم‌هایی تنها در یک بازه‌ی کوتاه زمانی قابل پذیرش بوده و در بازه‌های زمانی طولانی مدت که دنیای واقعی را ترسیم می‌نمایند نتایج خوبی را از خود ارائه نمی‌دهند.

در میان ابعاد زمینه‌ای موجود، بعد زمان؛ یعنی خصیصه‌های زمینه‌ای مربوط به زمان مثل *time of the day, day of the week and season of the year* دارای این مزیت هستند که به راحتی قابل جمع آوری بوده و هر سیستمی تقریباً می‌تواند برچسب زمانی<sup>۷</sup> امتیازدهی/ مصرف آیتم را ذخیره کند. علاوه بر این بعد زمان می‌تواند به عنوان یک ورودی با ارزش برای بهبود کیفیت پیشنهادات به کار گرفته شود [۷].

سیستم‌های پیشنهاد دهنده‌ی مبتنی بر زمان می‌توانند به عنوان نوع خاصی از *CARS*<sup>۸</sup> در نظر گرفته شوند. مشخصه اصلی آن‌ها، استفاده از اطلاعات زمینه‌ای زمان در بعضی از گام‌های فرایند پیش‌بینی امتیاز می‌باشد. بر اساس *Merriam-Webster*، زمان به عنوان "یک زنجیره غیر فضایی که در واحد رویدادهایی که به دنبال یکدیگر از گذشته تا به حال و تا آینده قرار می‌گیرند، اندازه گرفته می‌شود" به دلیل این انعطاف پذیری در مفهوم زمان و واحد اندازه‌گیری، نمایش‌های مختلفی از اطلاعات زمینه‌ای

زمان می‌تواند استفاده شود. برای مثال، زمان ممکن است به عنوان مقداری پیوسته که مقادیرش زمان‌های خاصی که آیتم‌ها امتیازدهی/ مصرف شده‌اند مدلسازی شود. ( برای مثال یک برچسب زمانی مثل *January 1<sup>st</sup>, 200, at 00:00:00*) گزینه بعدی تعیین مقادیر دسته بندی شده با در نظر گرفتن دوره زمانی علاقه در حوزه‌ی مربوط به پیشنهادات می‌باشد. برای نمونه، در حوزه‌ی توریسم، مقدار فصلی مثل *season of the year={ hot\_season, cold season}* مناسب باشد، در حالی که در حوزه‌ی موسیقی یا فیلم، مقدار *time of the week={workday, weekend}* ممکن است قابل لمس‌تر باشد. یک مدلسازی سلسله مراتبی در صورت استفاده می‌تواند به کنترل درجه‌ی درشت بودن<sup>۹</sup> اطلاعات زمینه‌ای زمان کمک کند. ( برای مثال *day of the week={Monday, Tuesday, ..Saturday}*---->*time of the week*) در این مفهوم، باید متذکر شد که ذخیره سازی برچسب زمانی امتیازدهی از آنجایی که سبب استخراج نمایشات مختلف از زمینه زمان و شامل هر دو مقدار پیوسته و دسته بندی شده می‌شود قابل انعطاف‌ترین گزینه می‌باشد. به طور عام، مدل‌های پیشنهاد دهنده‌ی مبتنی بر زمان، اطلاعات زمانی جمع آوری شده را که مربوط به علایق صریح گذشته‌ی کاربر می‌باشد، به کار می‌گیرند. برای مثال برچسب‌های زمانی مربوط به امتیازات. هرچند سایر منابع اطلاعات زمانی نیز می‌تواند جمع آوری و به کار گرفته شود. مثل زمان مصرف/ خرید آیتم، زمان ورود آیتم به کاتالوگ سیستم و زمان رجیستر شدن کاربر به تعاملات سیستم. به‌طور خلاصه، کاربران ممکن است در بازه‌های زمانی متفاوت دارای علایق متفاوت بوده، بنابراین اگر

<sup>۷</sup> Timestamp

<sup>۸</sup> Context Aware Recommender System

<sup>۹</sup> Granularity

اطلاعات مربوط به زمان انتساب برچسب توسط کاربر را در روند ارائه پیشنهاد توسط سیستم مورد استفاده قرار دهیم، خود سبب افزایش قابل ملاحظه ای در صحت پیشنهادات خروجی به کاربر می‌شویم. یکی از چالش‌های موجود در سیستم‌های پیشنهاد دهنده منبع، ترکیب موثر و سودمند اطلاعات به دست آمده از زمان با برچسب‌های انتسابی به منابع جهت افزایش صحت خروجی پیشنهادات است.

در این مقاله، ما اثر همزمان ارتباط میان منبع، برچسب، کاربر و زمان انتساب برچسب را به صورت مدل ۴ بعدی به کمک تنسور در نظر گرفته‌ایم. همچنین در این مقاله با استفاده از ویژگی تجزیه‌ی موجود در تنسورها روابط پنهان و مفید میان داده کشف و مورد استفاده قرار گرفته است. ارزیابی سیستم پیشنهادی بر روی داده‌های دنیای واقعی در مقایسه با سایر مقالات، نشان‌دهنده بهبود نتایج و همچنین افزایش صحت پیشنهادات در ارائه منابع به کاربر می‌باشد.

بخش‌های بعدی مقاله به صورت زیر سازماندهی شده است: در بخش ۲ کارهای مربوطه تشریح گردیده است، در بخش ۳ مفاهیم مربوط به عملیات تنسور قرار گرفته است، در بخش ۴ سیستم پیشنهادی معرفی شده است، در بخش ۵ نتایج مرتبط با پیاده‌سازی و در بخش ۶ نتیجه‌گیری قرار گرفته است.

## کارهای مرتبط

در این بخش تعدادی از کارهای مرتبط را که به دو دسته اصلی تقسیم کرده‌ایم بیان می‌کنیم: (۱) سیستم‌های پیشنهاد دهنده‌ی سنتی (۲) سیستم‌های پیشنهاد دهنده مبتنی بر زمینه

سیستم‌های پیشنهاد دهنده سنتی

سیستم‌های پیشنهاد دهنده‌ی سنتی [۹، ۱۰] از اطلاعات عمومی کاربر، آیتم برای ارائه پیشنهاد استفاده می‌کنند و به سه گروه اصلی تقسیم می‌شوند؛ مبتنی بر مشارکت (از اطلاعات کاربران مشابه کاربر جاری استفاده می‌کنند)؛ مبتنی بر محتوا (تنها مشخصات محتوایی آیتم مورد بررسی قرار می‌گیرد)؛ ترکیبی (به صورت ترکیبی از مشخصه‌های مبتنی بر مشارکت و محتوا به منظور حل مشکل "شروع سرد" [۱۱] عمل می‌کند). در سال‌های اخیر همزمان با ورود سیستم‌های برچسب‌زنی، برچسب‌ها در روش فیلترینگ مشارکتی جهت ارائه پیشنهاد به طور گسترده مورد استفاده قرار گرفته‌اند.

در واقع گسترش روزافزون سیستم‌های برچسب‌زنی اجتماعی، بستری جهت ارائه پیشنهادات شخصی-سازي شده فراهم نموده است. در سال‌های اخیر فعالیت‌های محققین به‌طور چشمگیری بر روی خصوصیات برچسب‌های ساخته شده‌ی کاربر، مانند واژگان مبهم، مترادف و تجزیه و تحلیل برچسب‌ها متمرکز شده است [۱۲]. اخیراً کارهای زیادی برای کشف دانش بر پایه برچسب‌ها صورت پذیرفته است که منجر به ساخت مدل کاربر [۱۳-۱۵]، پیش بینی برچسب [۱۶، ۱۷]، جستجوی وب معنایی [۱۸]، استخراج شبکه‌های اجتماعی [۱۹] شده است.

روش مشارکتی ارائه شده در [۲۰] با استفاده از ویژگی‌های تنسور، رابطه‌ی ۳ تایی میان کاربر، منبع و برچسب را به صورت همزمان در نظر گرفته است و با تجزیه‌ی تنسور ایجاد شده از چنین اطلاعاتی، رابطه‌ی پنهان میان منبع و برچسب کشف شده است. از این رو منبع به دست آمده از چنین تجزیه‌ای جهت ارائه پیشنهاد به کاربر مورد استفاده قرار گرفته و سبب بهبود کیفیت در نتایج پیشنهادات سیستم شده است. در این روش هرچند تأثیر همزمان اطلاعات برچسب، کاربر و منبع در نظر گرفته شده ولی تأثیر گذر زمان را که باعث افزایش کیفیت پیشنهادات می‌شود، در نظر نگرفته‌اند در صورتی که با در نظر گرفتن گذر زمان، پیشنهادات ارائه شده برای مسئله گذر زمان و تغییر علائق کاربران در دنیای واقعی مناسب‌تر خواهند بود.

در [۲۱] نیز یک روش فیلترینگ مشارکتی مبتنی بر برچسب ارائه شده است که در ابتدا مجموعه برچسب‌های کاندید هر کاربر را درون یک مدل قرار داده و سپس با استفاده از شبکه بیزین اقدام به ارائه پیشنهاد منبع به کاربر مورد نظر نموده است.

در [۲۲] یک روش عمومی برای دخالت دادن برچسب‌های انتسابی توسط کاربران به الگوریتم مرسوم دو حالته‌ی پیشنهاد دهنده‌ی (کاربر-آیتم) ارائه شده است. به طوری که ارتباط ۳ گانه‌ی بین کاربر، برچسب و آیتم را به ۳ ارتباط دوتایی از طریق ماتریس‌های (کاربر-آیتم)، (کاربر-برچسب) و (آیتم-برچسب) شکستند. سپس یک روش ترکیبی برای اتصال همبستگی بین این عناصر به کار گرفتند. این روش به دلیل در نظر نگرفتن ارتباط همزمان و ۳ گانه‌ی بین کاربر، برچسب و آیتم، از کارایی خوبی برای ارائه پیشنهادات برخوردار نمی‌باشد.

یک روش مشارکتی مبتنی بر برچسب در [۲۳] ارائه شده که شباهت کاربران را بر اساس فاصله‌ی معنایی میان برچسب‌های انتسابی توسط آن‌ها محاسبه می‌کند.

۲-۲ سیستم‌های پیشنهاد دهنده‌ی مبتنی بر زمینه

سیستم‌های پیشنهاد دهنده‌ی مبتنی بر زمینه [۲۴] به عنوان موضوعات تحقیقاتی مهمی در سال‌های اخیر پدیدار شدند [۲۵] که به عنوان نمونه می‌توان به روش ارائه شده در [۲۶] برای پیشنهاد فیلم و همچنین پیشنهاد موسیقی در [۷] اشاره نمود. روش‌های پیشنهاد دهنده‌ی مبتنی بر زمینه به سه گروه اصلی تقسیم می‌شوند: پیش‌فیلترینگ<sup>۱۰</sup>، پس‌فیلترینگ<sup>۱۱</sup> و مدلسازی زمینه‌ای<sup>۱۲</sup> [۱].

پیش‌فیلترینگ از اطلاعات زمینه‌ای برای انتخاب یا ساخت مرتبط‌ترین داده‌ی دوبعدی (کاربر × منبع) جهت تولید پیشنهادات استفاده می‌کند. در حقیقت در این روش زمینه‌ی مورد نظر به‌عنوان یک پرس‌وجو برای انتخاب داده‌های مرتبط استفاده می‌شود. پس

فیلترینگ از داده‌های زمینه‌ای در هنگام تولید پیشنهاد صرف نظر می‌کند و منابع نامربوط (در زمینه داده شده) را نادیده می‌گیرد یا امتیاز پیشنهاد را مطابق با زمینه، زمانی که لیست پیشنهادات آماده است تنظیم می‌کند، روش مدلسازی زمینه‌ای از اطلاعات زمینه‌ای مستقیماً در تابع پیشنهاد خود به‌عنوان پیش‌بینی-کننده‌ی صریح امتیازات کاربر در مورد منبع استفاده می‌کند. در [۸] انواع روش‌های مبتنی بر زمینه مورد مقایسه قرار گرفته‌اند. راه‌حل‌های مبتنی بر فاکتور-سازی تنسور که روش ما نیز شامل آن می‌شود، در گروه مدلسازی زمینه‌ای قرار می‌گیرد.

بعضی از روش‌های استاندارد مبتنی بر مدل، می‌توانند دارای چندین بعد باشند. یک مدل ارائه شده توسط Oku et al<sup>۱۳</sup>، SVM مبتنی بر زمینه [۲۷] می‌باشد. نویسندگان SVM را در یک فضای چند بعدی در نظر گرفته و ابرصفحات جداگانه‌ای را پیدا می‌کنند. نتایج آزمایشات آن‌ها نشان می‌دهد که پیشنهادات مبتنی بر زمینه نسبت به زمانی که از اطلاعات زمینه‌ای استفاده نمی‌کنند بهتر عمل می‌کند. روش دیگری که مسئله پیشنهادات مبتنی بر زمینه را از جنبه‌ی چند بعدی در نظر گرفته در [۲۴] ارائه شد که یک مدل چند بعدی مبتنی بر تکنیک پردازش تحلیلی بر خط<sup>۱۴</sup> برای پشتیبانی تصمیم‌گیری که هنوز مبتنی بر روش پیش‌فیلترینگ می‌باشد، معرفی می‌کنند.

در [۲۸] یک روش پیش‌فیلترینگ با تقسیم بندی پروفایل کاربر به چندین ریز پروفایل<sup>۱۵</sup> معرفی شده است که در آن تقسیم‌بندی با توجه به زمان و بر مبنای رخداد‌های صورت گرفته برای کاربر انجام شده است.

در [۲۹] یک روش مبتنی بر کاهش که در گروه روش پیش‌فیلترینگ قرار می‌گیرد برای مدل پیشنهادات چند بعدی که اطلاعات زمینه‌ای مثل زمان و مکان را در فرایند پیشنهادات قرار می‌دهد ارائه شده است.

<sup>۱۳</sup> Support Vector Machine

<sup>۱۴</sup> OLAP

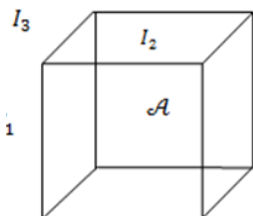
<sup>۱۶</sup> micro-profile

<sup>۱۰</sup> Pre-filtering

<sup>۱۱</sup> Post-filtering

<sup>۱۲</sup> Contextual modeling

استفاده قرار گرفته است [۳۰, ۳۱]. یک تنسور یک ماتریس چند بعدی است برای مثال یک تنسور  $N$  بعدی  $\mathcal{A}$ ، به صورت  $\mathcal{A} \in R_{I_1 \times \dots \times I_N}$  با عناصر  $a_{i_1, \dots, i_N}$  نشان داده می‌شود. در بخش‌های زیر تنسور با حروف بزرگ مربوط به دستنویس (مثل  $\mathcal{A}$  و  $\mathcal{B}$ )، ماتریس‌ها با حروف بزرگ (مثل  $A$  و  $B$ )، اعداد با حروف کوچک (مثل  $\alpha$  و  $b$ ) و بردارها با حروف کوچک توپر نشان داده می‌شوند (مثل  $a$  و  $b$ ). شکل ۱ نمایشی از تنسور ۳ بعدی  $\mathcal{A} \in R_{I_1 \times I_2 \times I_3}$  می‌باشد.



شکل ۱: تنسور ۳ بعدی  $\mathcal{A} \in R^{I_1 I_2 I_3}$

#### الف) تجزیه تنسور

مدل تاکر<sup>۱۸</sup> و مدل تحلیلی فاکتور موازی (PARAFAC) (که به عنوان تجزیه کانونی شناخته می‌شود) شناخته‌ترین و رایج‌ترین روش‌های تجزیه تنسور [۳۲] می‌باشد.

پارافک توسعه‌ای از مدل‌های فاکتورسازی دوسویه به داده‌های چندسویه می‌باشد و می‌تواند به عنوان تجزیه تنسور به صورت ترکیبی خطی از تنسورهای با رتبه-یک نمایش داده شود. مشخصه مفیدی از پارافک که موجب محبوبیت آن شده است، این است که راه حلی منحصر به فرد پیدا می‌کند به گونه‌ای که ماتریس‌های اجزا با جایگشت و مقیاس-پذیری ستون‌ها به طور منحصر به فرد تعیین می‌شوند. هرچند پارافک به عنوان مدل محدود در نظر

یک سیستم فیلترینگ مشارکتی مبتنی بر زمینه در [۲۸] ارائه شده است که علاقه‌ی کاربر را در موقعیت-های مختلف زمینه‌ای مورد بررسی و پیش بینی قرار می‌دهد؛ بنابراین کاربرانی که از نظر فکری در زمینه-های مورد نظر با کاربری خاص دارای ویژگی مشترکی باشند در یک گروه قرار گرفته و از اطلاعات چنین کاربرانی که به عنوان همسایه کاربر یاد می‌شود برای ارائه پیشنهادات موثر و مفید استفاده می‌شود. بعضی از کارهای انجام شده در زمینه‌ی فیلترینگ مشارکتی اطلاعات زمان را برای افزایش صحت در پیشنهادات دخالت داده‌اند.

در [۱۲] اطلاعات برچسب و زمان انتساب برچسب را با یکدیگر ترکیب نموده و در ارائه پیشنهاد استفاده نموده است، در این روش وزن زمانی منابع مربوط به کاربر و همچنین وزن برچسب‌های هر منبع را محاسبه و سپس با در نظر گرفتن پارامتری برای هر یک، آن‌ها را با یکدیگر ترکیب و وزن نهایی را به دست می‌آورد، در انتها با کشف کاربران همسایه اقدام به ارائه پیشنهاد به کاربر مورد نظر می‌نماید. در این روش به دلیل در نظر نگرفتن روابط همزمان بین عناصر موجود در سیستم‌های برچسب‌زنی، روابط پنهان موثر که منجر به افزایش عملکرد می‌شود، کشف نمی‌گردد و تنها از معیارهای شباهت عمومی برای کشف کاربران همسایه استفاده شده است.

در [۲۹] زمان خرید اجناس توسط کاربر و زمان ورود اقلام به بازار جهت بهبود صحت پیشنهادات در نظر گرفته شده است. در این مقاله از دو تابع امتیازدهی تکه‌ای<sup>۱۷</sup> جهت محاسبه وزن بر اساس اطلاعات موقتی استفاده شده است، سپس اطلاعاتی از جمله زمان ورود اقلام به بازار، زمان خرید کاربر، اختلاف زمانی مابین آن‌ها و ... مورد بررسی و تجزیه و تحلیل قرار گرفته‌اند. نتایج حاصل نشان می‌دهد که محاسبه و استفاده از چنین اطلاعاتی صحت پیشنهادات مبتنی بر فیلترینگ مشارکتی را افزایش می‌دهد.

#### تنسور و تجزیه‌ی تاکر

این بخش به طور مختصر راجع به تنسور و تجزیه‌ی آن توضیح می‌دهد. تحلیل چندخطی و جبر تنسور به طور موفقیت آمیزی در بسیاری از حوزه‌ها مورد

<sup>۱۸</sup> Tucker

<sup>۱۷</sup> piecewise

گرفته می‌شود، چراکه تنسور هسته محدود به قطری بودن می‌باشد.

تجزیه تاکر به عنوان مدل قابل انعطاف‌تر پارافک در نظر گرفته می‌شود و تعمیمی از عمل تجزیه مقادیر منحصر به فرد می‌باشد، به طوری که به طور موفقیت آمیزی در چندین حوزه مورد استفاده قرار گرفته است [۳۳, ۳۴] به همین دلیل مابقی مقاله بر روی تجزیه تاکر تمرکز می‌کند. جزئیات بیشتری راجع به روش‌های تجزیه تنسور عمومی در دو مقاله مروری [۳۵, ۳۲] می‌تواند یافت شود.

ب) تجزیه مقادیر منحصر به فرد و LSI

تجزیه مقادیر منحصر به فرد<sup>۱۹</sup> [۳۶] یک ماتریس  $F_{I_1 \times I_2}$  همانند عبارت زیر می‌تواند به صورت حاصل ضرب سه ماتریس نوشته شود:

$$F = \begin{pmatrix} & \\ & \\ & \end{pmatrix} \quad (۱)$$

به طوری که  $U$ ، ماتریسی با بردارهای منحصر به فرد چپ  $F$ ،  $V^T$  ترانهادهی ماتریس  $V$  که بردارهای منحصر به فرد راست  $F$  را شامل می‌شود و همچنین  $S$  که ماتریس قطری از مقادیر منحصر به فرد و مرتب شدهی  $F$  می‌باشد [۳۷].

ج) تجزیه تاکر

تجزیه تاکر<sup>۲۰</sup> برای تنسور، تجزیه مقادیر

منحصر به فرد را به ماتریس‌های چند بعدی تعمیم می‌دهد [۳۱]. برای مثال جهت اعمال این الگوریتم بر روی تنسور ۳ بعدی ابتدا ۳ ماتریس از تنسور به صورت زیر استخراج می‌شود:

$$A_1 \in \mathcal{R}^{I_1 \times I_2 \times I_3} \quad A_2 \in \mathcal{R}^{I_2 \times I_1 \times I_3} \quad A_3 \in \mathcal{R}^{I_1 \times I_2 \times I_3}$$

به  $A_1, A_2, A_3$  به ترتیب ماتریس‌های آشکار یک

وجهی<sup>۲۱</sup>، دو وجهی<sup>۲۲</sup> و سه وجهی<sup>۲۳</sup> از تنسور  $\mathcal{A}$  گفته می‌شود.

سپس ضرب  $n$  وجهی از تنسور  $N$  بعدی  $\mathcal{A} \in \mathbb{R}^{I_1 \times \dots \times I_N}$  با یک ماتریس  $U \in \mathbb{R}^{I_n \times I_n}$  به صورت  $\mathcal{A} \times_n U$  تعریف می‌شود. نتیجهی ضرب  $n$  وجهی، یک تنسور  $(I_1 \times I_2 \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N)$  می‌باشد که عناصر آن به صورت زیر تعریف می‌شود:

$$\begin{pmatrix} & \\ & \\ & \end{pmatrix} \quad (۲) \\ =$$

از آنجایی که تنسور چهار بعدی برای کار ما مدنظر می‌باشد؛  $n \in \{1, 2, 3, 4\}$  تنها ضرب‌های یک وجهی، دو وجهی، سه وجهی و چهار وجهی استفاده می‌شود.

با در نظر گرفتن ضرب  $n$  وجهی، تجزیه مقادیر منحصر به فرد بر روی یک ماتریس دو بعدی باقاعده (مثل تنسور دو بعدی) به صورت زیر می‌تواند نوشته شود [۳۱]:

$F = S \times_1 (U^{(1)}) \times_2 (U^{(2)})$   
به طوری که  $U^{(1)} = (u_1^{(1)} \ u_2^{(1)} \ \dots \ u_{I_1}^{(1)})$  یک ماتریس یکتا با ابعاد  $(I_1 \times I_1)$  می‌باشد،  
 $U^{(2)} = (u_1^{(2)} \ u_2^{(2)} \ \dots \ u_{I_2}^{(2)})$  یک ماتریس یکتا با ابعاد  $(I_2 \times I_2)$  و  $S$  یک ماتریس با ابعاد  $(I_1 \times I_2)$  و مشخصه زیر می‌باشد:

$$S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_{\min\{I_1, I_2\}}) \\ \text{نظم: } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min\{I_1, I_2\}} \geq 0$$

با تعمیم این شکل از عمل تجزیه مقادیر منحصر به فرد، تجزیه تاکر از تنسور ۳ بعدی  $\mathcal{A}$  می‌تواند به صورت زیر تعریف شود [۳۱]:

$$\hat{\mathcal{A}} = S \times_1 (U_{c_1}^{(1)}) \times_2 (U_{c_2}^{(2)}) \times_3 (U_{c_3}^{(3)}) \quad (۳)$$

به طوری که  $U^{(1)}, U^{(2)}, U^{(3)}$  شامل بردارهای نرمال (بردارهای یک وجهی - دو وجهی و سه وجهی

<sup>۱۹</sup> Singular Value Decomposition(SVD)

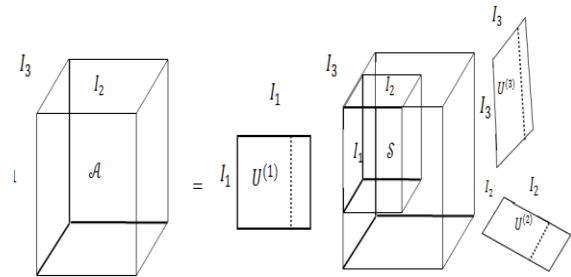
<sup>۲۰</sup> Tucker Decomposition

<sup>۲۱</sup> 1-mode matrix unfoldings

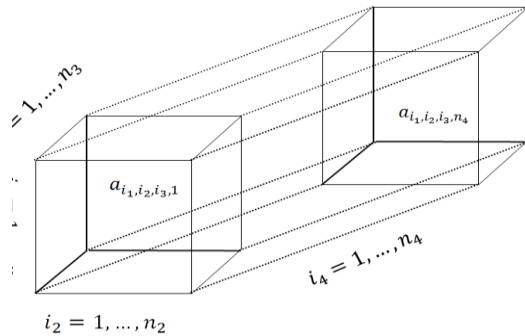
<sup>۲۲</sup> 2-mode matrix unfoldings

<sup>۲۳</sup> 3-mode matrix unfoldings

منحصر به فرد نامیده می‌شوند) و  $K$  تانسور هسته می‌باشد و خاصیت تعامدی<sup>۲۴</sup> دارد. شکل ۲ تجزیه تاکر را نشان می‌دهد.



گیرد. عناصر  $\mathcal{A}$  توسط رکورد  $\circ$  تایی  $\{u, t, l, tm, p\}$  می‌تواند نشان داده شود، به طوری که  $p$  احتمال علاقه‌ی کاربر  $u$  که برچسب  $t$  را بر روی آیتم  $i$  در ماه  $tm$  بزند را، اندازه می‌گیرد. بنابراین آیتم‌ها می‌توانند بر اساس وزن‌های مرتبط با  $\{u, t, tm\}$  به کاربر پیشنهاد شوند. شکل ۳ یک تانسور ۴ بعدی  $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$  را با عناصر  $a_{i_1, \dots, i_4}$  نشان می‌دهد:



## روش پیشنهادی

این بخش الگوریتمی به اسم HOSVD [31] را برای پیاده‌سازی تجزیه تاکر تانسور و چگونگی تولید پیشنهاد منبع با جزئیات توصیف می‌کند. برای درک مطلب در ابتدا یک طرح ریزی<sup>۲۵</sup> از روش ارائه شده از طریق یک مثال فراهم کرده و سپس گام‌های الگوریتم ارائه شده را تجزیه و تحلیل می‌کنیم.

به منظور تشریح اینکه چگونه روش ارائه شده کار می‌کند، الگوریتم را با مثالی دنبال می‌کنیم. فرض کنید در جدول ۱ اطلاعات مربوط به برچسب‌زنی سه کاربر را داریم.

جدول ۱: اطلاعات برچسب‌زنی کاربران

user	tag	item	time	weight
U <sub>ser1</sub>	Tag <sub>1</sub>	Item <sub>1</sub>	Month <sub>1</sub>	1
U <sub>ser1</sub>	Tag <sub>2</sub>	Item <sub>1</sub>	Month <sub>1</sub>	1
U <sub>ser1</sub>	Tag <sub>3</sub>	Item <sub>1</sub>	Month <sub>1</sub>	1
U <sub>ser2</sub>	Tag <sub>1</sub>	Item <sub>1</sub>	Month <sub>1</sub>	1
U <sub>ser3</sub>	Tag <sub>1</sub>	Item <sub>1</sub>	Month <sub>1</sub>	1
U <sub>ser3</sub>	Tag <sub>3</sub>	Item <sub>1</sub>	Month <sub>1</sub>	1
U <sub>ser1</sub>	Tag <sub>2</sub>	Item <sub>1</sub>	Month <sub>2</sub>	1
U <sub>ser2</sub>	Tag <sub>2</sub>	Item <sub>1</sub>	Month <sub>2</sub>	1
U <sub>ser3</sub>	Tag <sub>3</sub>	Item <sub>1</sub>	Month <sub>2</sub>	1
U <sub>ser2</sub>	Tag <sub>3</sub>	Item <sub>2</sub>	Month <sub>1</sub>	1
U <sub>ser1</sub>	Tag <sub>1</sub>	Item <sub>1</sub>	Month <sub>1</sub>	1

## ۱-۴ مثال

زمانی که از سیستم برچسب‌زنی مشارکتی استفاده می‌شود، جهت بازیابی راحت‌تر آیتمی در آینده، کاربر  $u$  برچسب  $t$  را به آیتم  $i$  انتساب می‌دهد. بنابراین بعد از مدتی سیستم برچسب‌زنی از مجموعه‌ای از داده برخوردار است. لازم به ذکر است که در هنگام برچسب‌زنی کاربر، زمان مربوط به عمل انتساب برچسب آن نیز ذخیره می‌شود. از آنجایی که گذر زمان بر روی علائق کاربران تاثیر به سزایی دارد بنابراین داده‌های موجود در سیستم-های برچسب‌زنی را می‌توان به صورت رکوردهای ۴ تایی  $\{u, t, i, tm\}$  در نظر گرفت. روش ارائه شده، HOSVD [31] را به تانسور ۴ بعدی ساخته شده از این داده‌های کاربردی اعمال می‌کند. پارامتر زمان به صورت ماه‌های سال و بدون همپوشانی با ماه سال-های متفاوت در نظر گرفته شده است. الگوریتم به عنوان ورودی داده‌های کاربردی را به شکل تانسور  $\mathcal{A}$  دریافت کرده و در خروجی، تانسور بازسازی شده  $\hat{\mathcal{A}}$  را می‌دهد. تانسور  $\hat{\mathcal{A}}$  ارتباط بین کاربر، برچسب، منبع و زمان انتساب برچسب را اندازه می-  


---

<sup>۲۴</sup> Orthogonality

<sup>۲۵</sup> Outline



User	Tag1	Ite	Mo	0.85
3		m3	nth2	38
User	Tag1	Ite	Mo	0.22
2		m2	nth2	15
User	Tag1	Ite	Mo	0.46
2		m3	nth2	08

همان‌طور که مشاهده می‌شود نتایج خروجی HOSVD بسیار جالب می‌باشد چرا که رابطه‌ی انجمنی جدیدی بر مبنای همبستگی پنهان کشف شده است. همان‌طور که نتایج نشان می‌دهند، وزن رکوردهایی که قبلاً دارای مقدار یک بوده‌اند، تغییر کرده است و حتی رکوردهای جدیدی با وزن‌های جدید نمایان شده‌اند، به طوری که قبل از ساخت تنسور دارای مقدار صفر بودند. فرض کنید که قرار است به کاربر ۲ در ماه دوم آیتمی پیشنهاد شود، آیتم‌هایی که با وزن جدید برای کاربر بدست آمده‌اند، آیتم ۳ و آیتم ۲ می‌باشند. در جدول ۲ رکورد  $\{u_2, t_1, i_2, month_2\}$  دارای وزن ۰.۴۶۰۸ و رکورد  $\{u_2, t_1, i_2, month_2\}$  دارای مقدار ۰.۲۲۱۵ شده است. با بررسی روابط موجود در اطلاعات برچسب-زنی درمی‌یابیم که عمل تجزیه به‌درستی صورت گرفته و حتی آیتم ۳ را که کاربر ۲ قبلاً برچسب نزده بود، از طریق روابط دوستان به‌درستی کشف کرده است. از آنجایی که در صورت وجود تعداد آیتم‌های زیاد موجود در مجموعه داده‌های برچسب زنی و تغییر وزن آن‌ها در تنسور بازسازی شده انتخاب مناسب‌ترین آیتم برای پیشنهاد باید صورت بگیرد ما بدین صورت عمل می‌کنیم که مثلاً برای ارائه پیشنهاد به کاربر ۲ مجموع وزن‌های برچسب‌های انتسابی به آیتم‌های ۲ و ۳ در ماه دوم را جداگانه جمع کرده و پس از مرتب سازی نزولی به کاربر پیشنهاد می‌دهیم. نتایج پیشنهاد قابل قبول می‌باشد چرا که آیتم‌های پیشنهادی با کشف روابط موجود پیشنهاد شده‌اند. روش ارائه شده قادر به کشف قوانین انجمنی پنهان

ser3	1	m2	nth2	
U	Tag	Ite	Mo	1
ser1	1	m3	nth2	
U	Tag	Ite	Mo	1
ser3	1	m3	nth2	

یک تنسور ۴ بعدی  $\mathcal{A} \in \mathbb{R}^{3 \times 3 \times 3 \times 2}$  از اطلاعات

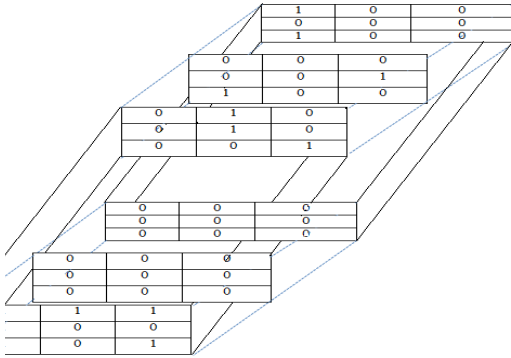
برچسب‌زنی موجود می‌تواند ساخته شود. تعداد دفعات وقوع (به عنوان وزن نشان داده شده است) تمام رکوردهای ۴ تایی (کاربر، برچسب، آیتم، زمان) مقادیر اولیه عناصر تنسور  $\mathcal{A}$  را تشکیل می‌دهند که در جدول ۱ نشان داده شده‌اند.

بعد از اعمال HOSVD (جزئیات مربوط به نحوه‌ی اعمال آن در بخش ۴-۲ بیان شده است) تنسور بازسازی شده‌ی  $\hat{\mathcal{A}}$  می‌تواند محاسبه شود که در جدول ۲ نشان داده شده است:

جدول ۲: اطلاعات برچسب زنی بعد از اعمال الگوریتم

user	tag	ite	time	weight
User	Tag1	Ite	Mo	1.15
1		m1	nth1	10
User	Tag2	Ite	Mo	0.47
1		m1	nth1	35
User	Tag3	Ite	Mo	0.91
1		m1	nth1	17
User	Tag1	Ite	Mo	0.58
2		m1	nth1	18
User	Tag1	Ite	Mo	1.06
3		m1	nth1	53
User	Tag3	Ite	Mo	0.84
3		m1	nth1	38
User	Tag2	Ite	Mo	0.71
1		m1	nth2	28
User	Tag2	Ite	Mo	0.36
2		m1	nth2	03
User	Tag3	Ite	Mo	0.40
3		m1	nth2	56
User	Tag3	Ite	Mo	0.12
2		m2	nth1	10
User	Tag1	Ite	Mo	0.40
3		m2	nth2	56
User	Tag1	Ite	Mo	0.91
1		m3	nth2	17

بر روی آیتم I در ماه tm زده است، ابتدا تنسور ۴ بعدی  $\mathcal{A} \in \mathbb{R}^{u \times t \times i \times tm}$  را می‌سازیم، به طوری که u تعداد تمام کاربران، I تعداد تمام آیتم‌ها، t مجموعه تمام برچسب‌ها و tm تعداد ماه‌های موجود در مجموعه داده می‌باشد. هر عنصر تنسور میزان علاقه‌ی کاربر u با برچسب t بر روی آیتم i و در ماه tm را نشان می‌دهد.



شکل ۵: تنسور اولیه برای مثال جاری

شکل ۵: تنسور اولیه برای مثال جاری  
 $\mathcal{A} \in \mathbb{R}^{1 \times \dots \times 1 \times N}$  می‌تواند به ماتریس تبدیل شود بدین صورت که تمام بردارهای ستونی / سطری یکی پس از دیگری قرار گیرد [20]. در روش ارائه شده از تنسور اولیه در هر چهار وجه ۲ ماتریس استخراج می‌شود. بنابراین بعد از استخراج تنسور  $\mathcal{A}$  برای هر ۴ وجه، چهار ماتریس  $A_{(1)}, A_{(2)}, A_{(3)}$  و  $A_{(4)}$  ساخته می‌شود.

### ۳) اعمال SVD بر روی هر وجه: تجزیه مقادیر

منحصر به فرد به چهار ماتریس استخراج شده‌ی  $A_{(i)}$ ،  $1 \leq i \leq 4$  اعمال می‌شود. این عمل منجر به تولید ماتریس‌هایی با شرایط زیر می‌شود:

$$A_i = U^{(i)} \cdot S^{(i)} \cdot (V^{(i)T}), \quad 1 \leq i \leq 4.$$

### ۴) محاسبه تخمین امتیاز پایین: در کاهش ابعاد

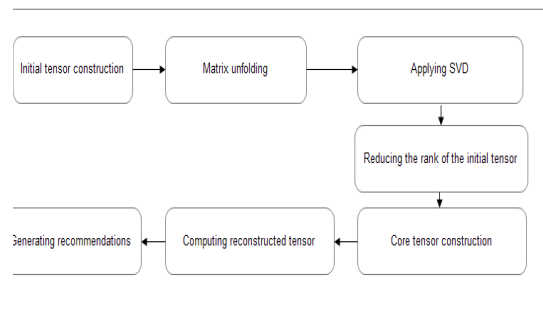
ماتریس‌ها، تخمین امتیاز پایین برای حذف مقادیر منحصر به فرد کوچک که نمایانگر نویز هستند، به کار می‌رود. به این عمل thin SVD گفته می‌شود و در

در بین کاربران، برچسب، آیتم و زمان انتساب برچسب می‌باشد، به طوری که قوانین پنهان منجر به تقویت روند پیشنهادات آیتم شده است.

### ۲-۴ HOSVD

این بخش اینکه چگونه HOSVD به تنسورها اعمال می‌شود و چگونه پیشنهاد آیتم‌ها بر طبق قوانین انجمنی نهان صورت می‌گیرد را تشریح می‌کند.

الگوریتم HOSVD اول یک تنسور  $\mathcal{A}$  بر اساس رکوردهای ۴ گانه  $\{u, t, I, tm\}$  از داده‌های برچسب‌زنی کاربران می‌سازد، هدف، استفاده از تمام ۴ عنصر کاربر، برچسب، آیتم و زمان انتساب برچسب‌ها می‌باشد که در سیستم‌های برچسب‌زنی از یکدیگر تاثیر می‌پذیرند. سپس عمل استخراج ماتریس از تنسور  $\mathcal{A}$  صورت می‌گیرد به طوری که ۴ ماتریس جدید استخراج می‌شوند. آنگاه عمل تجزیه مقادیر منحصر به فرد بر روی هر ماتریس اعمال شده و در نهایت تنسور هسته و تنسور بازسازی شده ساخته می‌شوند. از تنسور بازسازی شده برای پیشنهاد آیتم استفاده می‌شود، تمام این مراحل در ۷ گام و به صورت زیر خلاصه می‌شود:



شکل ۴: روند HOSVD

### ۱) ساخت تنسور اولیه $\mathcal{A}$ : از رکوردهای ۴ تایی

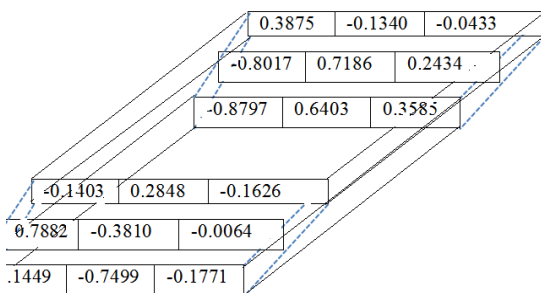
$\langle U, t, I, tm \rangle$  موجود در مجموعه داده‌های برچسب زنی که نشان می‌دهد کاربر u برچسب t را

$U^{(1)}, U^{(2)}, U^{(3)}, U^{(4)}$  انتخاب شدند، روش ارائه شده،

تفسیر هسته  $\mathcal{S}$  را به صورت زیر می‌سازد:

$$\mathcal{S} = \mathcal{A} \times_1 (U_{c_1}^{(1)})^T \times_2 (U_{c_2}^{(2)})^T \times_3 (U_{c_3}^{(3)})^T \times_4 (U_{c_4}^{(4)})^T$$
 به طوری که  $\mathcal{A}$  تفسیر اولیه،  $(U_{c_1}^{(1)})^T$  ترانهادهی ماتریس  $U^{(1)}$  کاهش یافته‌ی بعد  $c_1$ ،  $(U_{c_2}^{(2)})^T$  ترانهادهی ماتریس  $U^{(2)}$  کاهش یافته‌ی بعد  $c_2$ ،  $(U_{c_3}^{(3)})^T$  ترانهادهی ماتریس  $U^{(3)}$  کاهش یافته‌ی بعد  $c_3$  و  $(U_{c_4}^{(4)})^T$  ترانهادهی ماتریس  $U^{(4)}$  کاهش یافته‌ی بعد  $c_4$  می‌باشد.

شکل ۶ تفسیر هسته  $\mathcal{S}$  از مثال را نشان می‌دهد.

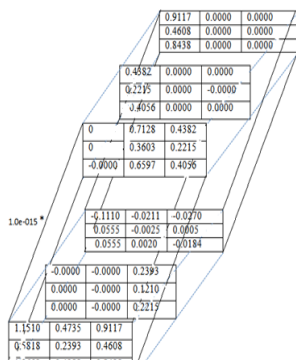


مثال جاری شکل ۶: تفسیر هسته

۶) ساخت تفسیر  $\hat{\mathcal{A}}$ : سرانجام تفسیر  $\hat{\mathcal{A}}$  با ضرب تفسیر هسته  $\mathcal{S}$  و ضرب وجهی (mode product) ۴ ماتریس  $U^{(i)}$  به صورت زیر بازسازی می‌شود:

$$\hat{\mathcal{A}} = \mathcal{S} \times_1 (U_{c_1}^{(1)}) \times_2 (U_{c_2}^{(2)}) \times_3 (U_{c_3}^{(3)}) \times_4 (U_{c_4}^{(4)})$$
 به طوری که  $\mathcal{S}$  تفسیر هسته کاهش یافته‌ی  $c_1, c_2, c_3, c_4$  ماتریس  $U^{(1)}$  کاهش یافته‌ی بعد  $c_1$ ،  $(U_{c_2}^{(2)})$  ماتریس  $U^{(2)}$  کاهش یافته‌ی بعد  $c_2$ ،  $(U_{c_3}^{(3)})$  ماتریس  $U^{(3)}$  کاهش یافته‌ی بعد  $c_3$  و  $(U_{c_4}^{(4)})$  ماتریس  $U^{(4)}$  کاهش یافته‌ی بعد  $c_4$  می‌باشد.

برای مثال جاری تفسیر بازسازی شده  $\hat{\mathcal{A}}$  در شکل ۷ نشان داده شده است.



LSI [۳۸] استفاده می‌شود. در سیستم پیشنهادی جهت کاهش تاثیر نویز اقدام به کاهش رتبه تفسیر اولیه نموده که این خود سبب کاهش میزان تنک بودن  $\mathcal{S}$  در ماتریس مورد نظر می‌شود. این عمل با تقریب تفسیر  $\mathcal{A}$  به یک تفسیر با رتبه پایین‌تر به دست می‌آید. در مورد کاهش ابعاد تفسیر، رتبه‌ی  $c_1, c_2, c_3, c_4$  باید محاسبه شود. به طوری که  $c_i$  ابعاد حفظ شده برای وجه  $i$  می‌باشد. برای محاسبه تخمین رتبه  $c_1, c_2, c_3, c_4$ ، مقادیر منحصر به فرد  $c_i$  و بردارهای واحد چپ متناظر با آن از  $U^{(i)}$  زمانی که SVD به ماتریس‌های  $A_i$  استخراج شده از هر وجه  $i$  اعمال می‌شود، نگه داشته می‌شوند. انتخاب مقادیر  $c_1, c_2, c_3, c_4$  ابعاد نهایی تفسیر هسته  $\mathcal{S}$  را تعیین می‌کند. از آنجایی که هر چهار ماتریس منحصر به فرد قطری  $S^{(1)}, S^{(2)}, S^{(3)}, S^{(4)}$  با اعمال SVD بر روی ماتریس‌های  $A_1, A_2, A_3, A_4$  محاسبه می‌شود، مقدار متفاوت  $c_i$  برای هر  $U^{(i)}$  ( $1 \leq i \leq 4$ ) استفاده می‌شود.

تعیین پارامترهای  $c_1, c_2, c_3, c_4$  در مدل‌های تاکر (مثل تجزیه مقادیر منحصر به فرد مرتبه‌ی بالا) عمل بسیار سختی می‌باشد [32]. یک گزینه‌ی تجربی استفاده از رتبه‌های نشان داده شده بوسیله‌ی تجزیه مقادیر منحصر به فرد بر روی داده‌های استخراج شده در هر وجه می‌باشد. این روش در چندین کار مورد استفاده قرار گرفته است [۳۰، ۳۹]. ما نیز در این مقاله از چنین روشی استفاده کرده‌ایم. در مثال  $c_1$  برابر با 3 و  $c_2=3, c_3=3, c_4=2$  می‌باشد.

۵) ساخت تفسیر هسته: تفسیر هسته  $\mathcal{S}$  ارتباطات متقابل بین کاربر، منبع، برچسب و زمان انتساب را نشان می‌دهد. زمانی که ابعاد ماتریس‌های

تمام مراحل کار ما در الگوریتم ۱ می‌تواند خلاصه شود.

Input:  
List of quadruples  
< users, tags, items, times >;  
Dimensions of users, tags, items and times  
 $c_1, c_2, c_3, c_4$ ;

Output:  
Generation of Item Recommendations

1: Initialization: From the quadruple (users, items, tag and time), we construct tensor  $\mathcal{A} \in \mathbb{R}^{|U| \times |t| \times |i| \times |tm|}$ , where  $|U|, |t|, |i|, |tm|$  are the number of users, items, tags and months respectively.

2: Calculate the matrix unfolding  $A_{(1)}, A_{(2)}, A_{(3)}$  and  $A_{(4)}$  from tensor  $\mathcal{A}$ .

3. Application of SVD on Each Mode: SVD is applied to the each matrix unfoldings  $A_{(i)}$

$1 \leq i \leq 4$ . This results to the following matrices:

$$A_i = U^{(i)} \cdot S^{(i)} \cdot (V^{(i)})^T, \quad 1 \leq i \leq 4.$$

4. Remove the least significant rows  $|U| - C_1, |i| - C_2, |t| - C_3, |tm| - C_4$  from  $U^{(1)}, U^{(2)}, U^{(3)}$  and  $U^{(4)}$  Respectively. Denote the result as  $U_{c_1}^{(1)}, U_{c_2}^{(2)}, U_{c_3}^{(3)}$ , and  $U_{c_4}^{(4)}$

5. Calculate the approximate core tensor  $\mathcal{S}$  as follows:

$$\mathcal{S} = \mathcal{A} \times_1 (U_{c_1}^{(1)})^T \times_2 (U_{c_2}^{(2)})^T \times_3 (U_{c_3}^{(3)})^T \times_4 (U_{c_4}^{(4)})^T$$

6. Approximate the original tensor by:

$$\hat{\mathcal{A}} = \mathcal{S} \times_1 U_{c_1}^{(1)} \times_2 U_{c_2}^{(2)} \times_3 U_{c_3}^{(3)} \times_4 U_{c_4}^{(4)}$$

7. Generation of Item Recommendations: The reconstructed tensor  $\hat{\mathcal{A}}$  measures the associations among the users, tags, and items at tm month. for recommending Let  $T_u$  be the set of tags that a user  $u$  has used to tag items which he/she like. For each item  $i$ , its total weight is computed as:

$$w_i = \sum_{t \in T_u} \bar{A}(u, t, i, tm)$$

After sorting items according to their  $w_i$ , the top  $N$  items with the highest weights are returned.

مثال جاری  $\hat{\mathcal{A}}$  شکل ۷: تنسور بازسازی شده

(۷) تولید پیشنهاد آیتم: تنسور  $\hat{\mathcal{A}}$  ارتباط بین کاربر، برچسب، منبع و زمان انتساب برچسب را اندازه می‌گیرد. عناصر  $\hat{\mathcal{A}}$  توسط رکورد ۵ تایی  $\{u, t, i, tm, p\}$  می‌تواند نشان داده شود، به طوری که  $p$  احتمال علاقه‌ی کاربر  $u$  که برچسب  $t$  را بر روی آیتم  $i$  در ماه  $tm$  بزند را اندازه می‌گیرد. بنابراین آیتم‌ها می‌توانند بر اساس وزن‌های مرتبط با  $\{u, t, tm\}$  به کاربر پیشنهاد شوند. مقادیر درونی تنسور بعد از مرحله بازسازی به دلیل تجزیه تنسور و روش‌های اعمال شده بر روی تنسور برای هر درایه  $\mathcal{A}(U, t, I, tm)$  تغییر می‌نماید.

پارامتر  $T_u$  را به عنوان مجموعه برچسب‌هایی که کاربر  $u$  بر روی عناصر مورد علاقه‌ی خود مورد انتساب قرار داده است در نظر بگیرید. از این رو جهت ارائه عناصر مورد علاقه‌ی کاربر می‌توان از پارامتر  $T_u$  استفاده نمود.

به ازای هر آیتم  $i$  مجموع وزن زیر را محاسبه می‌کنیم:

۱

سپس آیتم‌ها را مطابق با وزن  $w_i$  محاسبه شده‌ی آنها مرتب سازی نزولی نموده و از میان آن‌ها، تعداد  $N$  تا از آیتم‌هایی که بالاترین وزن را داشته در مجموعه پیشنهادی قرار داده و به کاربر  $u$  پیشنهاد می‌دهیم.

## نتایج پیاده‌سازی

این بخش عملکرد سیستم پیشنهادی را که به-

عنوان 4TB(4-order tensor based

recommendation) شناخته می‌شود با سیستم [۲۰]

که الگوریتم مبتنی بر برچسب برای پیشنهاد آیت می-

باشد مقایسه می‌کند. برای پیاده‌سازی 4TB جعبه

ابزار تنسور [۳۰] استفاده شده است که انجام تجزیه

تنسور را به صورت کارآمد با صرفه‌جویی در

مصرف حافظه و زمان اجرا، امکان‌پذیر می‌کند.

علاوه بر این، سیستم پیشنهادی در مقایسه با سیستم

[۱۲] که با در نظر گرفتن برچسب و زمان انتساب

آن‌ها توسط کاربران آیت‌های مناسب را پیشنهاد می-

دهد مورد بررسی قرار گرفته است.

دیتاست

دیتاست مورد استفاده، پایگاه داده واقعی

citeulike می‌باشد. citeulike وبسایتی برای محققان

جهت به اشتراک گذاری مراجع علمی از طریق انتساب

برچسب‌های شخصی به مقالات می‌باشد. ما از

دیتاستی که متعلق به بازه‌ی ۱۵ نوامبر ۲۰۰۴-

۱۳ فوریه ۲۰۰۷ می‌باشد استفاده کرده‌ایم. ۲۶۵۷۲۲۷

رکورد ۴ تایی به شکل کاربر-مقاله- برچسب- زمان

وجود دارد. این رکوردها به ۱۸۴۶۷ کاربر، ۵۵۷۱۰۱

مقاله، ۱۶۶۵۰۴ برچسب و ۳۲ ماه تعلق دارد. پس از

انجام عملیات پیش‌پردازش و اعمال p-core با مقدار

۳۰ بر روی دیتاست تعداد ۱۲۶۴۷۵ رکورد اطلاعاتی

باقی می‌ماند. پارامتر p-core تضمین کننده رخداد هر

کاربر، آیت و برچسب در حداقل p حاشیه نویسی

می‌باشد [۴۰].

پارامترها

برای الگوریتم 4TB ارائه شده پارامترهای زیر

مورد آزمون قرار گرفته‌اند. برای تعداد بردارهای

منحصر به فرد چپ باقی‌مانده (پارامترهای

c1,c2,c3,c4) مقادیر در بازه‌ی ۲۰٪-۸۰٪ در رابطه با

تعداد اصلی بردارهای منحصر به فرد، مورد آزمون

قرار گرفته‌اند. با افزایش مقادیر c1,c2,c3,c4 عملکرد

سیستم بهبود پیدا می‌کند. بنابراین مقدار درصد پیش-

فرض برای تمام پارامترهای c1,c2,c3 و c4 به ۸۰٪

تنظیم شده است. برای سیستم [۲۰] نیز مقادیر

پارامترهای c1,c2,c3 در بازه‌ی ۲۰٪-۸۰٪ مورد

آزمایش قرار گرفته و مقدار درصد پیش فرض برای

تمام پارامترهای c1,c2,c3 نیز به ۸۰٪ تنظیم شده

است.

برای سیستم [۱۲]، پارامتر  $\lambda$  از ۰ تا ۱ با فواصل

۰.۲ و پارامتر همسایگی k از ۵ تا ۱۰۰ با فواصل ۵

مورد تغییر قرار گرفتند. بهترین مقدار  $\lambda$  در بازه‌ی

۰.۴-۰.۹ می‌باشد، ما مقدار ۰.۸ را به عنوان پیش

فرض قرار دادیم. و بهترین مقدار k در بازه‌ی ۳۰-

۱۰۰ می‌باشد و ما مقدار ۳۰ را به عنوان مقدار پیش

فرض قرار دادیم.

پروتکل ارزیابی

پروتکل زیر برای ارزیابی مورد استفاده قرار

گرفته است: جهت انجام آزمایشات، ۵۰ درصد از

کاربران را به صورت تصادفی با احتمال انتخاب

یکنواخت انتخاب نموده و همچنین به ازای هر کاربر

۸۰ درصد از داده‌های اخیر را به عنوان مجموعه

داده‌ی آموزشی و ۲۰ درصد باقی مانده را به عنوان

مجموعه داده‌ی تست در نظر می‌گیریم. پارامترهای

ارزیابی precision، recall و f-measure [۴۱، ۴۲] جهت

نمایش کیفیت سیستم پیشنهادی اندازه‌گیری شده‌اند.

Recall یک متریک عمومی است که برای ارزیابی

سودمند بودن<sup>۲۹</sup> الگوریتم پیشنهادی به کار می‌رود و

احتمال به دست آوردن پیشنهادی درست را نشان

می‌دهد و به صورت معادله زیر عمل می‌نماید.

$$R = \frac{TP}{TP + FN} \quad (۴)$$

<sup>۲۹</sup> Utility

در این معادله پارامتر  $R_p$  مجموعه اقلام انتسابی توسط کاربر می باشد و  $R_r$  مجموعه اقلام پیشنهادی توسط سیستم می باشد.

Precision یک متریک عمومی است که برای اندازه گیری مفید بودن<sup>۳۰</sup> الگوریتم پیشنهادی به کار می رود و به صورت معادله زیر محاسبه می شود:

$$P = \frac{R_p}{R_p + R_r} \quad (5)$$

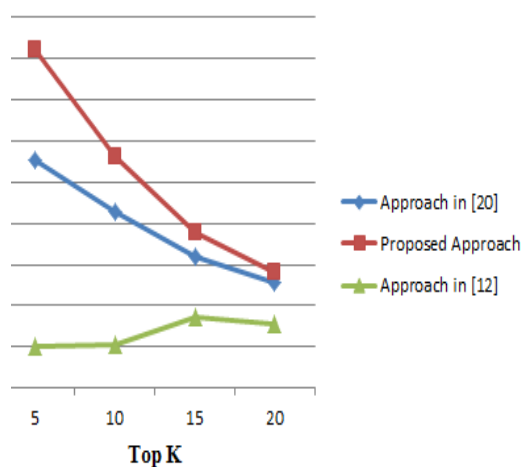
یکی دیگر از متریک های ارزیابی F-Measure می باشد که با استفاده از پارامترهای Recall(R) و Precision(P) به صورت معادله زیر به دست می آید:

$$F = \frac{2PR}{P+R} \quad (6)$$

از این متریک برای نمایش ارتباط بین Recall و Precision استفاده می شود.

#### نتایج

مقادیر recall, precision و f-measure برای هر سه سیستم 4TB پیشنهادی، سیستم [۲۰] و سیستم مبتنی بر برچسب و زمان [۱۲] و به ازای تعداد آیتم های پیشنهادی متفاوت در شکل های ۸، ۹ و ۱۰ نشان داده شده است.



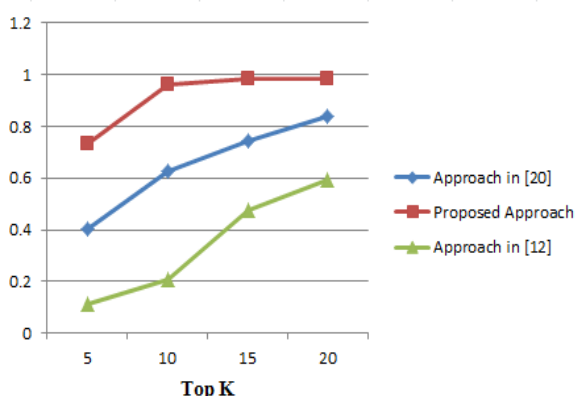
<sup>۳۰</sup> Usefulness

روش پیشنهادی و روش های مورد مقایسه Precision شکل ۸: مقایسه

در شکل ۸ محور عمودی معرف precision و محور افقی نمایانگر تعداد منابع پیشنهادی توسط سیستم می باشد. تعداد منابع پیشنهادی به کاربران در سیستم مورد بررسی، تغییر داده شده است که این خود نمایانگر مسئله درخواست  $k$  بالاترین<sup>۳۱</sup> می باشد

[۴۳]. با توجه به این مسئله کاربر قادر می باشد تعداد مرتبطترین منابعی که سیستم پیشنهاد می نماید را مشخص و دیگر نیازی به درگیر کردن کاربر برای جستجوی منابع مضاعف پیشنهاد شده نمی باشد [۴۳].

همان طور که در شکل ۸ مشاهده می شود میزان صحت سیستم پیشنهادی به ازای تمام مقادیر پیشنهاد شده بالاتر از روش های مورد مقایسه [۲۰, ۱۲] می باشد، همچنین سیستم پیشنهادی در مجموعه منابعی با تعداد کم نیز بسیار موثرتر از سیستم های مورد مقایسه عمل نموده است. این مزیت سبب می شود که سیستم در مجموعه منابعی با تعداد عناصر کم نیز دارای بالاترین صحت ممکن از لحاظ محاسبات آماری باشد و خود سبب راحتی کاربر در انتخاب عنصر مورد نظر با بالاترین احتمال ممکن از مجموعه داده های کم و بدون نیاز به جستجوی زیاد در مجموعه داده های نمایش داده شده می شود. از طرفی همان طور که در نمودار مشاهده می شود با افزایش تعداد پیشنهادات میزان صحت سیستم پیشنهادی پایین می آید به این دلیل است که تعداد منابع برچسب زده شده ی تست کاربران کمتر از تعداد پیشنهادات می باشد [۴۳]. به طور خلاصه علت افزایش صحت پیشنهادات در روش پیشنهادی به دلیل کشف روابط پنهان از طریق مدل سازی صحیح تنسور و اعمال همبستگی صحیح بین عناصر سیستم های برچسب زنی اجتماعی با در نظر گرفتن گذر زمان بر روی علایق کاربران می باشد. در حالی که در روش ارائه شده در [20] گذر زمان در ارائه پیشنهادات دخالت داده نشده است و تنها از اطلاعات عمومی موجود در سیستم های برچسب زنی استفاده کرده است. همچنین سیستم ارائه شده در [12] تنها از معیارهای شباهت عمومی برای کشف همسایگان کاربر و روابط دوگانه ی بین موجودیت های سیستم برچسب زنی استفاده کرده است.



تعداد منبع پیشنهادی ۵ (کمترین تعداد پیشنهاد) مقدار پارامتر F-Measure در روش پیشنهادی ۰.۷۷۵۹۰۵، در سیستم [۲۰] ۰.۴۶۷۹۴۷ و در سیستم [۱۲] ۰.۱۰۶۱۱ می‌باشد، نتایج آزمایشات نشان دهنده برتری سیستم پیشنهادی به نسبت سیستم‌های مورد مقایسه می‌باشد که این خود به دلیل استفاده از روابط چهار گانه در ساخت مدل، روابط انجمنی پنهان با در نظر گرفتن گذر زمان و اعمال همبستگی صحیح بین عناصر سیستم‌های برچسب زنی اجتماعی می‌باشد. همچنین سیستم پیشنهادی با انتخاب صحیح پارامترهای تجزیه تنسور روابط پنهانی را که سبب بهبود محاسبه‌ی معیارهای شباهت عمومی می‌شوند را کشف می‌نماید. در صورتی که سیستم [۲۰] تنها با در نظر گرفتن روابط همزمان بین کاربر، برچسب و منبع پیشنهادی را ارائه می‌دهد و تاثیر گذر زمان در تغییر علائق کاربران را مدنظر قرار نداده است و همچنین سیستم [12] با در نظر گرفتن روابط دوگانه‌ی موجودیت‌ها و معیارهای شباهت عمومی اقدام به ارائه پیشنهادت کرده است.

همان‌طور که نتایج آزمایشات نشان داده‌اند، سیستم پیشنهادی از کیفیت بالاتری در ارائه پیشنهادت در مقایسه با سیستم [۲۰، ۱۲] برخوردار است. از جمله دلایلی که باعث افزایش کیفیت سیستم پیشنهادی شده است می‌توان به کشف روابط پنهان موجود در عناصر تنسور اشاره نمود، چرا که با ساخت صحیح تنسور و همچنین عمل موثر تجزیه‌ی مقادیر، قادر به کشف وزن‌هایی از عناصر درون تنسور با در نظر گرفتن تاثیر همزمان گذر زمان بر روی علائق کاربران بوده که سبب افزایش صحت پیشنهادت در مجموعه نهایی و لیست ارائه شده به کاربر می‌باشد.

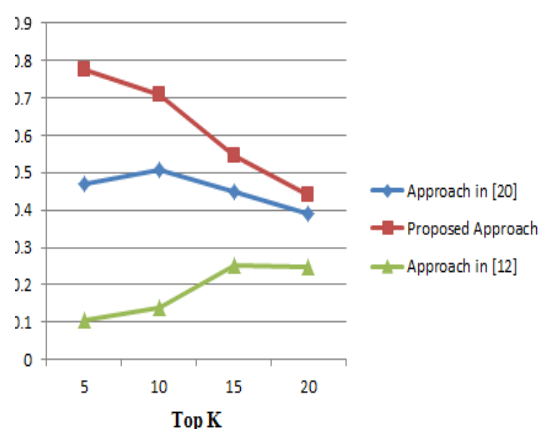
#### ۴-۵ بحث

ارائه پیشنهاد شامل دو بخش می‌باشد:

روش پیشنهادی و روش‌های مورد مقایسه Recall شکل ۹: مقایسه

در شکل ۹ محور عمودی نمایانگر Recall و محور افقی نشان دهنده‌ی تعداد منابع پیشنهادی توسط سیستم می‌باشد. همان‌طور که مشاهده می‌شود سیستم پیشنهادی نسبت به سیستم مورد مقایسه دارای رفتار بهتری در مجموعه‌های آماری و همچنین تعداد عناصر مجموعه‌های پیشنهادی جهت ارائه پیشنهادت می‌باشد همان‌طور که قبلاً ذکر شد این به دلیل کشف روابط پنهان از طریق مدل‌سازی صحیح تنسور با در نظر گرفتن گذر زمان بر روی علائق کاربران می‌باشد.

برای اینکه تاثیر precision و recall را با وزن هارمونیک نشان دهیم از معیار F-Measure استفاده نموده ایم. در شکل ۱۰ پارامتر ارزیابی F-Measure نشان داده شده است.



روش پیشنهادی و روش‌های مورد مقایسه F-Measure شکل ۱۰: مقایسه

همان‌طور که در شکل ۱۰ مشاهده می‌شود سیستم پیشنهادی به نسبت سیستم مورد مقایسه رفتار بهتری را از خود نشان می‌دهد. به‌طوری‌که در

بخشی که مربوط به محاسبات آفلاین برای تجزیه تنسور می‌باشد؛ ۲) بخشی که مربوط به محاسبات آنلاین برای بازیابی و مرتب سازی وزن‌ها از تنسور بازسازی شده می‌باشد. لازم به ذکر است که بخش آنلاین که بر روی تحمل کاربر تأثیر می‌گذارد در زمان واقعی (میلی ثانیه) می‌تواند اجرا شود. صرف نظر از بخش آفلاین، در سیستم‌های کاربردی واقعی (با تعداد زیاد کاربران، برچسب‌ها و آیتم)، مشخصه اصلی اسپارسیتی است، بدین معنا که اکثر مقادیر ورودی تنسور صفر هستند. بنابراین محاسبه‌ی بردارهای منحصر به فرد از ماتریس‌های استخراج شده‌ی بزرگ اسپارس و همچنین محاسبه ضرب یک تنسور اسپارس در یک مجموعه ای از ماتریس‌های متراکم (برای تنسور بازسازی شده) مشکل می‌شود. مقیاس پذیری این پروسه در [۳۰] مورد بررسی قرار گرفته است. به طوری که یک پیاده سازی از چارچوبی که سرعت محاسبات را افزایش می‌دهد و حافظه موجود را به طور بهینه مصرف می‌کند ارائه می‌دهد. با این روش تنسورهای خیلی بزرگ تنها با سخت افزارهای محدود می‌توانند ذخیره شوند.

برای برنامه های کاربردی دنیای واقعی استفاده از معماری موازی می‌تواند مقیاس پذیری روش [۳۰] را بهبود بخشد. سرانجام با نظر به بروز رسانی تدریجی زمانی که رکورد جدیدی اضافه می‌شود، روش folding-in ارائه شده در [۴۴] می‌تواند برای بروز رسانی SVD مورد استفاده قرار بگیرد. بنابراین بروز رسانی تدریجی از تجزیه تنسور (بخش آفلاین) به منظور جلوگیری از محاسبات دوباره‌ی آن از ابتدا قابل استفاده می‌باشد.

## نتیجه گیری و کارهای آینده

سیستم‌های برچسب‌زنی اجتماعی حاوی اطلاعات بسیار مفیدی می‌باشند، بنابراین ارائه سرویس‌های شخصی سازی شده‌ی مبتنی بر چنین اطلاعاتی از اهمیت بالایی برخوردار است. تحقیقات صورت گرفته بر روی چنین سیستم‌هایی در زمینه‌ی پیشنهاد آیتم پاسخگوی نیاز کاربر به طور کامل و در نظر گرفتن تمامی جنبه‌های مدل های سیستم‌های برچسب زنی اجتماعی نمی‌باشد.

در این مقاله یک سیستم جدید پیشنهاد دهنده‌ی منبعی معرفی نموده‌ایم که از اطلاعات برچسب و زمان انتساب آن‌ها جهت ارائه پیشنهادات استفاده می‌کند. همچنین نشان دادیم که تحلیل منطقی روابط چهار گانه منجر به تولید پیشنهاداتی با دقت بالاتر می‌شود.

سیستم پیشنهادی اطلاعات سیستم‌های برچسب-زنی اجتماعی را با تنسوری چهار بعدی جهت دریافت روابط چهار گانه‌ی بین کاربر، منبع، برچسب و زمان انتساب مدلسازی می‌کند. تحلیلات چند حالتی جهت کشف ارتباطات منطقی پنهان که منجر به افزایش کیفیت پیشنهادات می‌شود به کار گرفته شده‌اند. نتایج ارزیابی سیستم بر روی پایگاه داده‌ی دنیای واقعی و در مقایسه با سیستم ارائه شده در [۱۲] و [۲۰] نشان می‌دهد که تحلیل منطقی چهار گانه از کیفیت بالاتری در ارائه پیشنهادات برخوردار است و منجر به کشف روابط پنهان موجود در بین کاربران نیز می‌شود. همچنین از مزایای سیستم پیشنهادی می‌توان به استفاده از روابط چهار گانه در ساخت مدل، روابط انجمنی پنهان با در نظر گرفتن زمان و اعمال همبستگی صحیح بین عناصر سیستم‌های برچسب زنی اجتماعی اشاره نمود. علاوه بر این سیستم پیشنهادی با انتخاب صحیح پارامترهای تجزیه تنسور روابط پنهانی را که سبب بهبود محاسبه‌ی معیارهای شباهت عمومی می‌شوند را کشف می‌نماید.

سیستم پیشنهادی را می‌توان به صورتی سیستمی برای پیشنهاد کاربر و همچنین برچسب، علاوه بر پیشنهاد آیتم به کاربران جهت کارهای آینده مورد بررسی و توسعه قرار داد.

## مراجع

[1] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in *Recommender*



recommender systems," *ACM Transactions on Information Systems (TOIS)*, vol. 22, pp. 5-53, 2004.

[12] N. Zheng and Q. Li, "A recommender system based on tag and time information for social tagging systems," *Expert Systems with Applications*, vol. 38, pp. 4575-4587, 2011.

[13] C. M. A. Yeung, N. Gibbins, and N. Shadbolt, "A study of user profile generation from folksonomies," presented at the Social Web and Knowledge Management Workshop, Beijing, China, 2008.

[14] Z. Yun and F. Boqin, "Tag-based user modeling using formal concept analysis," presented at the 8th IEEE International Conference on Computer and Information Technology, 2008.

[15] E. Michlmayr and S. Cayzer, "Learning user profiles from tagging data and leveraging them for personal(ized) information access," in *Workshop on Tagging and Metadata for Social Information Organization, 16th International World Wide Web Conference*, 20.07

[16] B. Sigurbjornsson and R. V. Zwol, "Flickr tag recommendation based on collective knowledge," presented at the 17th international conference on World Wide Web, 2008.

[17] Z. Xu, Y. Fu, J. Mao, and D. Su, "Towards the semantic web: Collaborative tag suggestions," in *15th international conference on World WideWeb*, Edinburgh, Scotland, 2006.

[18] P. Heymann, & Koutrika, G. , "Can social bookmarking improve web search?," presented at the international conference on Web search and web data mining (WSD ),(New York, USA.

.2008 ,

[19] X. Li, L. Guo, and Y. E. Zhao, "Tag-based social interest discovery," presented at the 17th international conference on World Wide Web, New York, USA, 2008.

[20] A. Nanopoulos, "Item Recommendation in Collaborative Tagging Systems," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 41, 2011.

[21] H. Kim, Ji, A., & Jo, G. , "Collaborative filtering based on collaborative tagging for

*Systems Handbook.*, F. Ricci, Rokach, L., Shapira, B., Kantor, P.B, Ed., ed, 2011.

[2] T. Hussein, T. Linder, G. Werner, and J. Ziegler, "Hybreed: A software framework for developing context aware hybrid recommender systems," *User Modeling and User-Adapted Interaction*, 2014.

[3] A. K. Dey, "Understanding and using context," *Personal and Ubiquitous Computing*, vol. 5, pp. 4-7, 2001.

[4] L. Baltrunas, "Context-aware collaborative filtering recommender systems," PhD Dissertation, Free University of Bozen-Bolzano, 2011.

[5] S. S. Anand and B. Mobasher, "Contextual recommendation: From Web to Social Web: Discovering and Deploying User and Content Profiles," *Berlin: Springer*.2007 ,

[6] S. Berkovsky, T. Kuflik, L. Aroyo, D. Heckmann, A. Kroner, F. Ricci, *et al.*, "Predicting user experiences through cross-context reasoning," in *14th Workshop on Adaptivity and User Modeling in Interactive Systems*, 2006, pp. 27-31.

[7] L. Baltrunas and X. Amatriain, "Towards time-dependant recommendation based on implicit feedback," in *Workshop on Context-Aware Recommender Systems*, New York, NY, 2009.

[8] U. Panniello, A. Tuzhilin, M. Gorgoglione, C. Palmisano, and A. Pedone, "Experimental comparison of pre-vs. post-filtering approaches in context-aware recommender systems," presented at the Third ACM Conference on Recommender Systems, 2009.

[9] J. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, 1998, pp. 43-52.

[10] G. Karypis, "Evaluation of item-based top-n recommendation algorithms," in *tenth international conference on Information and knowledge management (CIKM'01)*, 2001.

[11] J. Herlocker, Konstan, J., Terveen, L., & Riedl, J. , "Evaluating collaborative filtering

- [31] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, pp. 12.2000, 1278–53
- [32] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, pp. 6-20, 2009.
- [33] J. Sun, D. Shen, H. Zeng, Q. Yang, Y. Lu, and Z. Chen, "Cubesvd: A novel approach to personalized Web search," in *Proceedings of the 14th international conference on World Wide Web*, 2005, pp. 382–390.
- [34] H. Wang and N. Ahuja, "A tensor approximation approach to dimensionality reduction," *International Journal of Computer Vision*, vol. 76, pp. 217–229, 2008.
- [35] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, pp. 455–500, 2009.
- [36] M. Berry, S. Dumais, and G. O'Brien, "Using linear algebra for intelligent information retrieval," *SIAM Review*, vol. 37, pp. 573–595, 1995.
- [37] A. Loizou and S. Dasmahapatra, "Recommender systems for the semantic Web," in *Proceedings of ECAI*, 2006, pp. 76–81.
- [38] G. Furnas, Deerwester, S., & Dumais, S., "Information retrieval using a singular value decomposition model of latent semantic structure," in *11th annual international ACM SIGIR conference on Research and development in information retrieval*, 1988.
- [39] J. Sun, D. Tao, and C. Faloutsos, "Beyond streams and graphs: Dynamic tensor analysis," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2006, pp. 374–383.
- [40] R. Jaschke, Marinho, L., Hotho, A., Schmidt-Thieme, L., & Stumme, G., "Tag Recommendations in Folksonomies," in *11th European conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)*. 2007.
- [41] P. Symeonidis, A. Nanopoulos, and Y. Manolopoulos, "A Unified Framework for Providing Recommendations in Social Tagging enhancing the quality of recommendation," *Electronic Commerce Research and Applications*, vol. 9, pp. 73-83, 2010.
- [22] K. Tso-Sutter, B. Marinho, and L. Schmidt-Thieme, "Tag-aware recommender systems by fusion of collaborative filtering algorithms " in *Proceedings of the 2008 ACM symposium on Applied computing*, 20, 08pp. 1995–1999.
- [23] S. Zhao, N. Du, A. Nauerz, X. Zhang, Q. Yuan, and R. Fu, "Improved recommendation based on collaborative tagging behaviors," in *Proceedings of the 2008 ACM conference on recommender systems (RecSys'08)*, 2008, pp. 413–416.
- [24] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin, "Incorporating contextual information in recommender systems using a multidimensional approach," *ACM Transactions on Information Systems (TOIS)*, vol. 23 pp. 103-145, 2005.
- [25] A. Said, S. Berkovsky, and E. W. De Luca, "Putting things in context: Challenge on context-aware movie recommendation," presented at the Workshop on Context-Aware Movie Recommendation, Barcelona, Spain, 2010.
- [26] T. Bogers, "Movie recommendation using random walks over the contextual graph," in *2nd Workshop on Context-Aware Recommender Systems*, Barcelona, Spain, 2010.
- [27] K. Oku, S. Nakajima, J. Miyazaki, and S. Uemura, "Context-aware svm for context-dependent information recommendation.," in *Proceedings of the 7th international Conference on Mobile Data Management*, 2006.
- [28] A. Chen, "Context-aware collaborative filtering system: Predicting the user's preference in the ubiquitous computing environment," *Location-and Context-Awareness*, pp. 244-253, 2005.
- [29] T. Q. Lee, Y. Park, and Y. T. Park, "A time-based approach to effective recommender systems using implicit feedback," *Expert systems with applications*, vol. 34, pp. 3055–3062, 2008.
- [30] T. G. Kolda and J. Sun, "Scalable tensor decompositions for multi-aspect data mining," in *Proceeding of the 8th IEEE International Conference on Data Mining(ICDM)*, 2008.

Systems Based on Ternary Semantic Analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, pp. 179-192, 2010.

[42] I. H. Witten and E. Frank. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (second ed.).

[43] M. N. Jelassi, S. B. Yahia, and E. M. Nguifo, "A personalized recommender system based on users' information in folksonomies," in *Proceedings of 22nd international conference on World Wide Web companion*, Geneva, Switzerland, 2013.

[44] B. Sarwar, J. Konstan, and J. Riedl, "Incremental singular value decomposition algorithms for highly scalable recommender systems," in *Proceedings of 5th International Conference on Computer and Information Science*, 2002, pp. 27–28.